Strengthened Lindblad inequality: applications in non equilibrium thermodynamics and quantum information theory.

A.E. Allahverdyan, D.B. Saakian

Yerevan Physics Institute

Alikhanian Brothers St.2, Yerevan 375036, Armenia

A strengthened Lindblad inequality has been proved. We have applied this result for proving a generalized H-theorem in non equilibrium thermodynamics. Information processing also can be considered as some thermodynamic process. From this point of view we have proved a strengthened data processing inequality in quantum information theory.

There are close connections between statistical thermodynamics and information theory [2] [5]. It is well known that physical ideas played an important role as sources of information theory [2]. On the other hand, the concept of information is crucial for understanding some important physical problems such as Maxwells "demon" [5] or the general problem of quantum correlation between two subsystems [10].

In this paper we concentrate on the two connected problems. These are the H-theorem problem in non equilibrium quantum statistical thermodynamics and the problem of quantum data processing in quantum information theory. The concepts of entropy (or other entropy-like measures) and H-theorem are particularly important in *quantum* statistical physics because a correct definition is only possible in the framework of quantum mechanics. In classical theory entropy can only be introduced in a somewhat limited and artificial manner [4] [7].

Suppose that a quantum system is described by a density matrix $\rho(t)$ at the moment t. In the general case evolution of the non equilibrium system in *markoffian* regime is described by some general quantum evolution operator

$$\hat{K}(t',t)\rho(t) = \rho(t') \tag{1}$$

In the most general case $\hat{K}$ must be linear, completely positive and trace-preserving [11] [15], and has standard representation

$$\hat{K}\rho = \sum_{\mu} A_{\mu}^{\dagger}\rho A_{\mu}, \quad \sum_{\mu} A_{\mu}A_{\mu}^{\dagger} = \hat{1}. \tag{2}$$

which was introduced in [15]. This representation is equivalent to the so called unitary representation where the non-unitary evolution of the system is regarded as a part of the unitary evolution of some larger system. Eq. (2) contains unitary transformations, nonselective measurements, partial traces, et all. For a non-markoffian case $\hat{K}$ also depends on the "history" from some initial time t_0 to t' [3].

If the evolution of the system is in the *stationary markoffian* regime then

$$\hat{K}(t',t) = \hat{K}(t'-t) \tag{3}$$

Stationary markoffian regime is a reasonable conjecture if the system is not far from equilibrium [1] [3], or if it can be described by a non-hermitian time-independent hamiltonian [3] or by quantum Langevin equations [19].

One of the most important quantities which can be defined for statistical systems is entropy [1] [3] [2] [5] [4] [7]. This quantity was introduced in quantum statistical physics by J. von Neumann.

$$S(\rho) = -\mathrm{tr}\rho \ln \rho \tag{4}$$

The concept of entropy has at least three main ingredients. In the first, entropy of a macroscopic statistical system can only increase if the system tends to an equilibrium. Therefore entropy is maximal at this state. This is the well known H-theorem. We want to stress that increasing of (4) with time can be violated if the evolution of the system is not stationary or markoffian [1] [3] or the system is open or mesoscopic [9]. For example, entropy can exhibit exactly periodic behavior for some open system [20]. This means that (4) is not the relevant statistical function for such systems. Entropy is additive function, and also invariant of an unitary transformation.

In the second, entropy can be regarded as a measure of the lack of information about a system. Therefore $S(\rho)$ should increase after coarse-graining procedure [2] [3] [7]

$$S(\sum_i p_i \rho_i) \geq \sum_i p_i S(\rho_i), \quad \sum_i p_i = 1, \quad p_i \geq 0 \tag{5}$$

Where the information about coordinate i is lost (i can be also continuous).

In the third entropy can be considered as a measure of the amount of chaos, or, to what extent the density matrix ρ can be considered as "mixed". Indeed, the non negative $S(\rho)$ is zero for a pure density matrix, and is maximal for homogeneous ρ. Eq (4) can also be viewed as one of the basic statements of equilibrium statistical physics [1] [3] [4]. For example, after several assumptions the most important relation in thermostatics: $TdS = dE + pdV$ can be derived from (4) (where all symbols have their ordinary meaning).

Now the following questions arise. Is it possible to define an entropy-like function for a mesoscopic statistical system or for an open system? Is it possible to save in this definition the main aspects of usual entropy? Large number of papers and books are devoted to these questions [1] [3] [6]. The answer is *"Yes"* at least in the case when the evolution of a system is stationary markoffian, and has a well defined stationary distribution. The concrete form of this distribution is not important. Before the definition we need some mathematics.

Quantum relative entropy between two density matrices ρ_1, ρ_2 is defined as follows

$$S(\rho_1||\rho_2) = \mathrm{tr}(\rho_1 \log \rho_1 - \rho_1 \log \rho_2). \tag{6}$$

This positive quantity was introduced by Umegaki [17] and characterizes the degree of 'closeness' of density matrices ρ_1, ρ_2. The properties of quantum relative information were reviewed by M.Ohya [16]. Here only two basic properties are mentioned.

$$S(\rho_1||\rho_2) \geq S(\hat{K}\rho_1||\hat{K}\rho_2). \tag{7}$$

$$S(\lambda\rho_1 + (1-\lambda)\rho_2||\lambda\sigma_1 + (1-\lambda)\sigma_2) \leq \lambda S(\rho_1||\sigma_1) + (1-\lambda)S(\rho_2||\sigma_2). \tag{8}$$

Where $0 \leq \lambda \leq 1$. The first inequality was proved by Lindblad [18].

Now for a system with stationary distribution ρ_{st}, and markoffian stationary evolution operator $\hat{K}$ the following function is defined

$$-S(\rho(t)||\rho_{st}) \tag{9}$$

This function is additive, and also increases after coarse-graining procedure as we see from (8). Further, eq. (7) which can be written as

$$-S(\hat{K}\rho(t)||\rho_{st}) \geq -S(\rho(t)||\rho_{st}) \tag{10}$$

is H-theorem for (9).

The definition (9) is closely related to the functions which are used in usual equilibrium statistical physics. A very large closed statistical system can be described by microcanonical distribution where ρ_{st} can be represented as an unit matrix (up to some unessential factors). In this case (9) reduces to eq. (4) (at least in the case of finite dimensional Hilbert space), and from (10) we have the usual H-theorem. Further it is well known that for a closed macroscopic system canonical and microcanonical distributions are equivalent (except some special cases like second-order phase transitions). But in some sense canonical distribution has larger area of application because it can describe some mesoscopic or quasi-open systems [3] [4]. If we take $\rho_{st} = \exp(-\beta H)/Z$ in eq. (9) (where β is inverse temperature, and H is hamiltonian) then

$$S(\rho(t)||\rho_{st}) - \ln Z = \mathrm{tr}(\rho \ln \rho) + \beta \mathrm{tr}(H\rho) = \beta F \tag{11}$$

Where F is usual free energy. Therefore for the case of canonical distribution we have a slightly different form of H-theorem: the free energy can only decrease if the system tends to equilibrium [1] [3] [4] [6].

Is the physically relevant generalization entropy is defined uniquely? This important question was investigated in [8]. The author showed that (7, 8) with some other mathematical conditions are sufficient for the determination of (6).

The conclusion is the following: (9) *is correct generalization of entropy to the more general case, and a generalized H-theorem can be proved with the assumptions about the evolution of the system only.*

Can we generalize (7) without any restrictions? If the answer is yes, then we can prove with this result a more general relation. Let us assume in formula (7) that

$$\hat{K} = c\hat{C}_1 + (1-c)\hat{C}_2, \tag{12}$$

where $\hat{C}_1$ is defined by kraussian representation $A_\mu = |\mu\rangle\langle 0|$, $\langle \mu|\acute{\mu}\rangle = \delta_{\mu\acute{\mu}}$, $\langle 0|0\rangle = 1$, $0 \leq c \leq 1$. In other words for any operator ρ: $\hat{C}_1\rho = |0\rangle\langle 0|$. Now from (7), (8) we get

$$\begin{aligned} S(\hat{K}\rho||\hat{K}\sigma) &= S(c\hat{C}_1\rho + (1-c)\hat{C}_2\rho||c\hat{C}_1\sigma + (1-c)\hat{C}_2\sigma) \\ &\leq cS(\hat{C}_1\rho||\hat{C}_1\sigma) + (1-c)S(\hat{C}_2\rho||\hat{C}_2\sigma) \leq (1-c)S(\rho||\sigma). \end{aligned} \tag{13}$$

We see that if $\hat{K}$ is represented in the form (12) the ordinary Lindblad inequality can be strengthened.

Now we need some general results from theory of linear operators [23]. Let two hermitian operators A and B have the spectrums $a_1 \leq ... \leq a_n$, $b_1 \leq ... \leq b_n$. For the spectrum $c_1 \leq ... \leq c_n$ of the operator $C = A + B$ we have

$$a_1 + b_k \leq c_k \leq b_k + a_n, \quad b_1 + a_k \leq c_k \leq a_k + b_n. \tag{14}$$

where $k = 1, ..., n$. If

$$
\begin{aligned}
\rho' = \hat{K}\rho = c\hat{C}_1\rho + (1-c)\hat{C}_2\rho \\
= c|0\rangle\langle 0| + (1-c)\sigma,
\end{aligned} \tag{15}
$$

and $\rho'_1 \leq ... \leq \rho'_n$, $\sigma_1 \leq ... \leq \sigma_n$ are the spectrums of ρ', σ then we have

$$
\begin{aligned}
&\rho'_1 - c \leq \sigma_1(1-c) \leq \min(\rho'_1, \rho'_n - c), \\
&\max(\rho'_1, \rho'_k - c) \leq \sigma_k(1-c) \leq \rho'_k,
\end{aligned} \tag{16}
$$

where $k = 2, ..., n$. We define $c(\hat{K}, \rho)$ as the minimal eigenvalue of ρ' and $c(\hat{K}) = \min_\rho c(\hat{K}, \rho)$ where minimization is taken by all density matrices for the fixed Hilbert space. With the well known results of operator theory [23] we can write

$$
c(\hat{K}) = \min_\rho \min_{\langle\psi|\psi\rangle=1} \langle\psi|\hat{K}\rho|\psi\rangle, \tag{17}
$$

where the second minimization is taken by all normal vectors in the Hilbert space. For any density matrix ρ we get to the formula (12) where c is defined in (17) and $\hat{C}_2$ is some general evolution operator. Now from (12,13,17) we get the strengthened Lindblad inequality

$$
(1-c)S(\rho_1||\rho_2) \geq S(\hat{K}\rho_1||\hat{K}\rho_2). \tag{18}
$$

The equations (17), (18) are our general results. Of course there are many evolution operators $\hat{K}$ with $c(\hat{K}) = 0$ but later we shall show that our results can be nontrivial because for some simple but physically important case $c(\hat{K})$ is nonzero. From (17), (18) we immediately get to the strengthened H-theorem which gives us some information about the speed of relative entropy decrease. An analog of (18) exists also in classical information theory [22]. Eq. (18) can be also regarded as a bound for entropy production. This quantity is very important in nonequilibrium statistical mechanics [1] [3].

Now about application of this result to quantum data processing.

Quantum information theory is a new field with potential applications for the conceptual foundation of quantum mechanics. It appears to be the basis for a proper understanding of the emerging fields of quantum computation, communication and cryptography (see [11] for references). Quantum information theory is concerned with quantum bits (qubits) rather than bits. Qubits can exist in superposition or entanglement states with other qubits, a notion completely inaccessible for classical mechanics. More general, quantum information theory contains two distinct types of problem. The first type describes transmission of classical information through a quantum channel (the channel can be noisy or noiseless). In such a scheme bits are encoded as some quantum states and only these states or their tensor products are transmitted. In the second case arbitrary superposition of these states or entanglement states are transmitted. In the first case the problems can be solved by methods of classical information theory, but in the second case new physical representations are needed.

Mutual information is the most important ingredient of information theory. In classical theory this quantity was introduced by C.Shannon [22]. The mutual information between two ensembles of random variables X, Y (for example these ensembles can be input and output for a noisy channel)

$$
I(X,Y) = H(Y) - H(Y/X), \tag{19}
$$

is the decrease of the entropy of X due to the knowledge about Y, and conversely with interchanging X and Y. Here $H(Y)$ and $H(Y/X)$ are Shannon entropy and mutual entropy [22].

Mutual information in the quantum case must take into account the specific character of the quantum information as it is described above. The first reasonable definition of this quantity was introduced by S.Lloyd [14], and independently by B.Schumacher and M.P.Nielsen [12]. Suppose a quantum system with density matrix

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|, \quad \sum_i p_i = 1. \tag{20}$$

We only assume that $\langle\psi_i|\psi_i\rangle = 1$ and the states may be nonorthogonal. The noisy quantum channel can be described by some general quantum evolution operator $\hat{K}$.

As follows from the definition of quantum information transmission, a possible distortion of entanglement of ρ must be taken into account. In other words a definition of mutual quantum information must contain the possible distortion of the relative phases of the quantum ensemble $\{|\psi_i\rangle\}$. Mutual quantum information is defined as [14] [12]

$$I(\rho;\hat{K}) = S(\hat{K}\rho) - S(\hat{1}^R \otimes \hat{K}(|\psi^R\rangle\langle\psi^R|)), \tag{21}$$

$$\hat{1}^R \otimes \hat{K}(|\psi^R\rangle\langle\psi^R|) = \sum_{i,j} \sqrt{p_i p_j} |\phi_i^R\rangle\langle\phi_j^R| \otimes \hat{K}(|\psi_i\rangle\langle\psi_j|). \tag{22}$$

Where $S(\rho)$ is the entropy of von Newman and ψ^R is a purification of ρ

$$|\psi^R\rangle = \sum_i \sqrt{p_i} |\psi_i\rangle \otimes |\phi_i^R\rangle, \quad \langle\phi_j^R|\phi_i^R\rangle = \delta_{ij}, \tag{23}$$

$$\mathrm{tr}_R |\psi^R\rangle\langle\psi^R| = \rho, \tag{24}$$

here $\{|\phi_i^R\rangle\}$ is some orthonormal set. The definition is independent of the concrete choice of this set [11]. Mutual quantum information is the decrease of entropy after the action of $\hat{K}$ due to the possible distortion of entanglement state. This quantity is not symmetric with respect to the interchanging of input and output and can be positive, negative or zero in contrast with Shannon mutual information in classical theory. It has been shown that (21) can be the upper bound of the capacity of a quantum channel [13]. Using this value the authors [13] have been proved the converse coding theorem for a quantum source with respect to the so called entanglement fidelity [11]. This fidelity is absolutely adequate for quantum data transmission or compression.

In the [12] the authors prove a data processing inequality

$$I(\rho;\hat{K}_1) \geq I(\rho;\hat{K}_2\hat{K}_1). \tag{25}$$

The quantum information can not increase after action of $\hat{K}$. In [13] we found an alternative derivation of this result which is simpler than the derivation of [12]. In the present paper we show that this inequality can be strengthened. The data processing inequality is a very important property of mutual information. This is an effective tool for proving general results and the first step toward identification of a physical quantity as mutual information.

Now we briefly recall the derivation of the data processing inequality. The formalism of relative quantum entropy is very useful in this context. We have

$$S(\hat{1}^R \otimes \hat{K}(|\psi^R\rangle\langle\psi^R|)||\hat{1}^R \otimes \hat{K}(\rho^R \otimes \rho))$$
$$= -S(\hat{1}^R \otimes \hat{K}(|\psi^R\rangle\langle\psi^R|)) + S(\rho^R) + S(\hat{K}\rho)). \tag{26}$$

Here

$$\rho^R = \sum_{i,j} \sqrt{p_i p_j} |\phi_i^R\rangle\langle\phi_j^R|\langle\psi_i|\psi_j\rangle. \tag{27}$$

Now from the Lindblad inequality we have

$$S(\hat{1^R} \otimes \hat{K}(|\psi^R\rangle\langle\psi^R|)||\hat{1^R} \otimes \hat{K}(\rho^R \otimes \rho))$$
$$\geq S(\hat{1^R} \otimes \hat{S}_1\hat{K}_2(|\psi^R\rangle\langle\psi^R|)||\hat{1^R} \otimes \hat{K}_1\hat{K}_2(\rho^R \otimes \rho)). \tag{28}$$

From this formula we have (25).

Now we can prove the strengthened data processing inequality. Let in (28) $\hat{K}_2$ is represented in the form (12). From (7,12) we get

$$S(\hat{1}^R \otimes \hat{K}_2\hat{K}_1(|\psi^R\rangle\langle\psi^R|)||\hat{1}^R \otimes \hat{K}_2\hat{K}_1(\rho^R \otimes \rho))$$
$$\leq -(1-c)S(\hat{1}^R \otimes \hat{C}_2\hat{K}_1(|\psi^R\rangle\langle\psi^R|)) + S(\rho^R) + (1-c)S(\hat{C}_2\hat{K}\rho)). \tag{29}$$

And we have

$$(1 - c(\hat{K}_2))I(\rho; \hat{K}_1) \geq I(\rho; \hat{K}_2\hat{K}_1). \tag{30}$$

Now we consider the simplest example of noisy quantum channel: Two dimensional, two- Pauli channel [21] with the following Krauss representation

$$A_1 = \sqrt{x}\hat{1}, \quad A_2 = \sqrt{(1-x)/2}\sigma_1, \quad A_3 = -i\sqrt{(1-x)/2}\sigma_2, \quad 0 \leq x \leq 1, \tag{31}$$

where $\hat{1}$, σ_1, σ_2 are the unit matrix and the first and the second Pauli matrices. The (31) also has physical meaning as an evolution operator for a two-dimensional open system. Any density matrix in two-dimensional Hilbert space can be represented in the Bloch form

$$\rho = (1 + \vec{a}\vec{\sigma})/2, \tag{32}$$

where $\vec{a}$ is a real vector with $|\vec{a}| \leq 1$. Now we have

$$\hat{K}_{TP}((1 + \vec{a}\vec{\sigma})/2) = (1 + \vec{b}\vec{\sigma})/2, \tag{33}$$

where $\vec{b} = (a_1 x, a_2 x, a_3(2x-1))$. After simple calculations we get

$$c(\hat{K}_{TP}) = (1 - |2x-1|)/2. \tag{34}$$

We conclude by reiterating the main results: the Lindblad inequality can be generalized. We have presented results not only about increasing of entropy and decreasing of mutual quantum information but also about the speed of these processes.

[1] F.Shlogl, Phys. Rep.,**62**, 268, (1980).

[2] R.L. Stratanovich, *Information Theory*, 1975, Moscow, Nauka.

[3] R.L. Stratanovich, *Nonequilibriun, Nonlinear Thermodynamics*, 1985, Moscow, Nauka.

[4] L.D. Landau, and E.M. Lifshits, Statistical Physics, (Moscow, Nauka, 1976).

[5] R.P. Poplavsky, *Thermodynamics of information processes*, 1981, Moscow, Nauka.

[6] The literature on this particular point is too broad to cite every work. A recent book of M.C. Mackey, *Time's Arrow: The Origins of Thermodynamic Behavior*, (Springer-Verlag, Berlin, 1992) provides a good review. See also: M.C. Mackey, Rev.Mod.Phys., **61**, 981, (1989); Yu.L. Klimontovich, Uspekhi Fizicheskikh Nauk, **164**, 811, (1994); Yu.L. Klimontovich, *Statistical Physics*, (Moscow, Nauka, 1982).

[7] A. Wehrl, Rev.Mod.Phys., **50**, 221, (1978).

[8] M.J. Donald, Comm.Math.Phys., **105**, 13, (1986).

[9] For a comprehensive review and further references about nonequilibrium steady states in open systems, see: B. Schmittmann and R.K.P. Zia, in *Phase Transitions and Critical Phenomena*, eds. C. Domb and J.L. Lebowitz (Academic Press, New York, 1996).

[10] S.M. Barnett and S.J.D. Phoenix, Phys. Rev. A **44**, 535, (1991).

[11] B.Schumacher, Phys. Rev. A **54**, 2614, (1996).

[12] B.Schumacher and M.A. Nielsen, Phys. Rev. A **54**, 2629, (1996).

[13] A.E. Allahverdyan and D.B. Saakian, eprint quant-ph/9702023.

[14] S. Lloyd, Phys. Rev. A **55**, R1613-1622, (1997); also e-print quant-ph/9604015.

[15] K.Kraus, Ann. Phys., **64**, 311, (1971).

[16] M.Ohya, Rep. Math. Phys., **27**, 19, (1989).

[17] H.Umegaki, Kodai Math. Sem. Rep., **14**, 59, (1962).

[18] G.Lindblad, Commun. Math. Phys., **40**, 147, (1975).

[19] M. Lax, Phys.Rev., **145**, 110, (1966).

[20] S.J.D. Phoenix and P.L. Knight, Ann. Phys.(N.Y) **186**, 381, (1988).

[21] C.H. Bennett, C.A. Fuchs, J.A. Smolin, eprint quant-ph/9611006.

[22] I.Csiszar, J.Korner, *Information Theory*, (Akademiai Kiado, Budapest,1981).

[23] F.Gantmacher, *The Theory of matrices*, Moscow, Nauka 1983.